Continuous Adjoint-Based Optimization of Hyperbolic Equations with Nonlinear Differential Equation Constraints on Periodic Boundary Conditions

Nhan T. Nguyen

***Abstract*— This paper presents a continuous adjoint-based optimization theory for a general closed-loop transport hyperbolic model controlled via a periodic boundary control to minimize a multi-objective cost functional. The periodic boundary control is subject to a nonlinear differential equation constraint, thus resulting in a coupling between the hyperbolic equation and the ordinary differential equation. Variational principles are used to derive the Pontryagin's minimum principle for optimality that results in a dual adjoint system. A numerical optimization method is implemented using the adjoint-based second-order gradient method to solve for the optimal trajectory of the control. Numerical methods for solving the hyperbolic equation using an explicit-scheme, wave splitting method and for solving the adjoint equation using an implicit scheme and a quasi-steady state method are described.**

I. INTRODUCTION

Hyperbolic partial differential equations (PDEs) are used to model transport systems whose information is carried from one point to another within those systems as a function of space and time [1]. Examples of transport systems are numerous such as fluid flow in gas distribution pipelines [2], air traffic systems [3], highway traffic systems [4], to name a few. These equations describe wave propagation that exists in transport systems to propagate information from one point to another within the continuum. As with any PDEs, boundary conditions are used to specify configurations of these transport systems. If the information is carried in one direction without returning to its starting position, then we say that the system is open-loop. An example of an open-loop system is gas flow through an aircraft engine. On the other hand, if the information returns to its starting position, then the system is said to be closed-loop. An example of a closed-loop system is the cardiovascular circulatory system. Boundary conditions associated with closed-loop systems are usually periodic in nature.

The flow of information is usually supplied at the system boundary by a forced process that provides a motive force to move the information along the way by wave propagation. For example, a common device for accomplishing this objective in fluid transport systems is a pump which supplies a positive pressure head to displace the fluid volume in the flow direction. Such a process whereby the control is applied at the boundary of the continuum is called a boundary control process. In real systems, boundary control processes are often governed by other auxiliary dynamical processes. For example, a positive-displacement pump may be driven by an electrical motor that imposes a constraint on the pump speed according to the motor torque dynamics. The pump speed in this case is a boundary control variable. Thus, in the present framework, we wish to study a boundary control process of a closed-loop transport system modeled by hyperbolic PDEs that is coupled to a dynamical system governed by nonlinear ordinary differential equations (ODEs) via a periodic boundary condition.

In this study, we will develop a continuous adjoint theory to deal with this type of closed-loop transport systems. The theory will then be applied to develop a numerical optimization for a closed-loop fluid transport system to minimize a multi-objective cost function. Solutions to hyperbolic adjoint systems are based on an implicit-scheme, first-order upwind method as well as a quasi-steady state method. We demonstrate the optimization method by a numerical example of a closed-loop fluid flow problem.

Nhan T. Nguyen is with NASA Ames Research Center, Intelligent Systems Division, M/S 269-1, Moffett Field, CA 94035

II. HYPERBOLIC TRANSPORT MODEL

Transport phenomena are governed by the conservation laws equations which dictate the conservation of some quantities such as traffic density, mass flux, and enthalpy. These equations are generally hyperbolic in nature. For a 1-D system, these equations are expressed in a conservation form as [5]

$$\frac{\partial \mathbf{y}}{\partial t} + \frac{\partial \mathbf{F}(\mathbf{y}, x)}{\partial x} + \mathbf{Q}(\mathbf{y}, x) = \mathbf{0} \qquad \forall x \in [0, L], t \in [0, T] \tag{1}$$

where $\mathbf{y}(x, t) : [0, L] \times [0, T] \rightarrow \mathbb{R}^n$ in class C^1 is a vector of conserved quantities, $\mathbf{F}(\mathbf{y}, x)$ is a flux function, and $\mathbf{Q}(\mathbf{y}, x)$ is a non-homogeneous source term. If $\mathbf{Q}(\mathbf{y}, x) = \mathbf{0}$, then the steady state solution of (1) is the conservation laws

$$\mathbf{F}(\mathbf{y}, x) = \mathbf{C} \qquad \forall x \in (0, L) \tag{2}$$

where $\mathbf{C}$ is a constant vector.

By explicit differentiation, (1) can be rewritten in as

$$\frac{\partial \mathbf{y}}{\partial t} + \mathbf{A}(\mathbf{y}, x)\frac{\partial \mathbf{y}}{\partial x} + \mathbf{B}(\mathbf{y}, x) = \mathbf{0} \qquad \forall x \in (0, L), t \in (0, T) \tag{3}$$

where $\mathbf{A}(\mathbf{y}, x) : \mathbb{R}^n \times [0, L] \rightarrow \mathbb{R}^n \times \mathbb{R}^n$ is a characteristic matrix such that $\mathbf{A}(\mathbf{y}, x) = \mathbf{F}_{\mathbf{y}}(\mathbf{y}, x)$ and $\mathbf{B}(\mathbf{y}, x) : \mathbb{R}^n \times [0, L] \rightarrow \mathbb{R}^n$ is a non-homogeneous source term such that $\mathbf{B}(\mathbf{y}, x) = \mathbf{Q}(\mathbf{y}, x) + \mathbf{F}_x(\mathbf{y}, x)$.

Equation (3) is a system of first order, quasilinear hyperbolic equations due to the fact that the matrix $\mathbf{A}$ has n real, distinct eigenvalues such that

$$\lambda_1(\mathbf{A}) < \lambda_2(\mathbf{A}) < \ldots \lambda_n(\mathbf{A})$$

for all $\mathbf{y}(x,t) \in \mathbb{R}^n$, $x \in [0, L]$, and $t \in [0, T]$.

Under this condition, the matrix $\mathbf{A}$ is diagonalizable using a similarity transformation [6]

$$\mathbf{A} = \mathbf{\Phi}\mathbf{\Lambda}\mathbf{\Phi}^{-1} \tag{4}$$

where $\mathbf{\Phi}$ is an matrix of the right eigenvectors and $\mathbf{\Lambda}$ is a diagonal matrix of the right eigenvalues of $\mathbf{A}$

$$\mathbf{\Lambda} = diag(\lambda_1, \lambda_2, \ldots, \lambda_m, \lambda_{m+1}, \lambda_{m+2}, \ldots, \lambda_n) \tag{5}$$

where $m < n$ is the number of negative eigenvalues.

The eigenvalues are the wave speeds of the transport systems and the direction of the wave propagation is called a characteristic direction. If $m > 0$, the information in the continuum is carried in both the upstream and downstream directions by negative wave speeds λ_i, $i = 1, 2, \ldots, m$ and positive wave speeds λ_i, $i = m+1, m+2, \ldots, n$; respectively. If the solution domain is $0 \le x \le L$, then for the information to be transported in the upstream direction by the negative wave speeds, information must exist at the boundary $x = L$. Similarly, information must also exist at the boundary $x = 0$ in order for the information to be carried downstream by the positive wave speeds. Therefore, the number of upstream and downstream boundary conditions must match the number of negative and positive eigenvalues. This is known as the boundary condition compatibility.

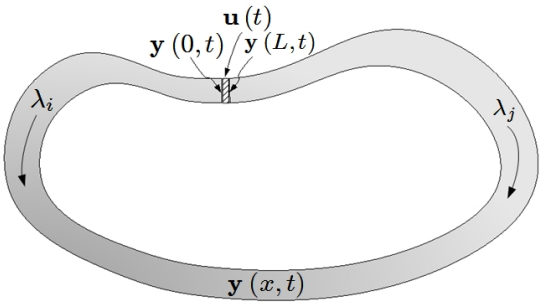

Fig. 1 - Closed-Loop Transport Model

In a closed-loop transport system, information is carried from one point to another and then returned back to the starting position as illustrated in Fig. 1. To enable this information flow, a periodic boundary control process is embedded within the system. For a closed-loop system, the boundary conditions at $x = 0$ are affected by the boundary conditions at $x = L$ since the information must be returned to its starting position. Thus, in general for a closed-loop system, we specify the following general nonlinear periodic boundary condition for (2)

$$\mathbf{y}(0,t) = \mathbf{g}(\mathbf{y}(L,t), \mathbf{u}(t)) \qquad \forall t \in [0, T] \tag{6}$$

where $\mathbf{u}(t) : [0, T] \to \mathbb{R}^m$ in class C^1 is a boundary control vector, and $\mathbf{g}(\mathbf{y}(L,t), \mathbf{u}) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is a nonlinear forcing function that relates the transport state vectors at $x = 0$ and $x = L$ and the boundary control vector $\mathbf{u}$.

Theorem 1: To ensure the boundary condition compatibility for all signs of eigenvalues, the Fréchet derivative or the Jacobian of $\mathbf{g}$ with respect to $\mathbf{y}(L,t)$ is required to have a full rank or

$$\dim\left[\mathbf{g}_{\mathbf{y}(L,t)}\right] = n \tag{7}$$

Proof: Let m be the number of negative eigenvalues. If $m = 0$, then from (6), there are n independent boundary conditions with $\mathbf{y}(0,t)$ that are sufficient to provide information for n positive eigenvalues. Thus, the compatibility for positive eigenvalues is satisfied. If $m = n$, since $\mathbf{g}_{\mathbf{y}(L,t)}$ is full rank from (7), then there are also n independent boundary conditions with $\mathbf{y}(L,t)$ that are sufficient to provide information for n negative eigenvalues. The compatibility for negative eigenvalues is thus satisfied. If $0 < m < n$, we choose $n - m$ independent boundary conditions with $\mathbf{y}(0,t)$. Then, From (7), there are m remaining boundary conditions with $\mathbf{y}(L,t)$. Since there are $m - n$ positive eigenvalues and m negative eigenvalues, then the compatibility for mixed-sign eigenvalues is thus satisfied. ■

In many real systems, boundary control processes are actually controlled by other auxiliary processes. These auxiliary processes may be dynamical so that their dynamics can be described by a nonlinear ODE

$$\dot{\mathbf{u}} = \mathbf{f}(\mathbf{y}(0,t), \mathbf{y}(L,t), \mathbf{u}, \mathbf{v}) \tag{8}$$

where $\mathbf{v}(t) : [0, T] \to \mathbb{R}^l$ is an auxiliary control vector that belongs to a convex subset of admissible auxiliary control $\mathcal{V}_{ad} \subseteq \mathbb{R}^l$, and $\mathbf{f}(\mathbf{y}(0,t), \mathbf{y}(L,t), \mathbf{u}, \mathbf{v}) : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^l \to \mathbb{R}^m$ is a nonlinear function.

Thus the auxiliary control vector $\mathbf{v}$ actually influences the boundary control vector $\mathbf{u}$, which in turn controls the behavior of the closed-loop transport system described by (3) and the periodic boundary condition (6).

III. CONTINUOUS ADJOINT OPTIMALITY

Optimal control and optimization theories of hyperbolic systems has been studied extensively in mathematical literature. Within the theoretical framework of systems governed by PDEs, control of such systems can exist as distributed control, boundary control, interior pointwise control, or others. Hou and Yan studied the long time behavior of solutions for an optimal distributed control problem for the Navier-Stokes equations [8]. Nguyen et al investigated a flow control problem with interior pointwise control [9]. Optimal control problems of transport systems with boundary control have been examined for many different types of constraints imposed on either state or control variables. Raymond and Zidani investigated necessary optimality conditions in the form of a Pontryagin's minimum principle for semilinear parabolic equations with pointwise state constraints and unbounded control [10]. Casas et al established second order sufficient conditions for local optimality of elliptic equations with pointwise constraints on the boundary control and equality and set-constraints on state variables [11]. Kazemi obtained adjoint equations for a degenerate hyperbolic equation [12].

Adjoint method is well-known in optimization theories as it provides an indirect method for solving optimization problems. For a transport system governed by hyperbolic equations, two types of adjoint formulation are used: discrete adjoint and continuous adjoint. Any hyperbolic equation can generally be discretized into a system of ODEs by means of numerical discretization techniques such as finite-difference or finite-element methods. If the adjoint method is formulated with the discretized hyperbolic equations, then this is known as discrete adjoint method. On the other hand, continuous adjoint method is normally applied directly to the original hyperbolic equations. This method has been used in aerodynamic design optimization studies involving the Euler and Navier-Stokes equations [13]. Nadarajah et al compared the discrete and continuous adjoint methods in aerodynamic design optimization and suggested that the continuous adjoint method affords a certain advantage over the discrete adjoint method for Navier-Stokes flow problems [14]. In the present study, we extend this method to the present hyperbolic system with nonlinear differential equation constraints on a periodic boundary condition. To formulate the continuous adjoint method for this system, we seek a solution of the hyperbolic system above that minimizes the following multi-objective cost functional

$$\min\ J(\mathbf{y},\mathbf{u},\mathbf{v}) = \int_0^T \int_0^L L_1(\mathbf{y})\,dxdt + \int_0^T L_2\left(\mathbf{y}^0,\mathbf{y}^L,\mathbf{u},\mathbf{v}\right)dt \quad (9)$$

where L_1 is an objective function defined over the continuum of the transport system, L_2 is an objective function defined on the system boundary, and the superscripts 0 and L are short hand notations denoting the associated vector quantity evaluated at $x=0$ and $x=L$, respectively.

The following assumptions are required:

(A1): Equation (3) admits smooth solutions for shock-free conditions.

(A2): $\mathbf{v}\in\mathcal{L}_2$, the space of real value functions in $\mathbb{R}^l$ for which the norm $\|\mathbf{v}\|$ is square-integrable.

(A3): The Fréchet derivatives of L_1, L_2, $\mathbf{B}$, $\mathbf{g}$, $\mathbf{f}$, and $\mathbf{h}$ with respect to $\mathbf{y}$, $\mathbf{y}^0$, $\mathbf{y}^L$, $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ exist and are bounded so as to satisfy the Lipschitz condition.

We note that (1) also has discontinuous solutions known as entropy solutions [7]which will not be treated here.

The transport system above is posed as a boundary control problem of hyperbolic equations with nonlinear differential equation constraints.

Lemma 1: Let D be a nonlinear differential operator and D^* be its adjoint differential operator such that for some $\mathbf{z}(x,t)\in\mathbb{R}^n$ and $\boldsymbol{\lambda}(x,t)\in\mathbb{R}^n$

$$D\mathbf{z} = \mathbf{A}\frac{\partial \mathbf{z}}{\partial x} \quad (10)$$

$$D^*\boldsymbol{\lambda} = \frac{\partial}{\partial x}\left(\mathbf{A}^\top\boldsymbol{\lambda}\right) \quad (11)$$

where the superscript $\top$ is the transpose operator, then the following inner product operation in $\mathcal{L}_2$ is equivalent

$$\langle D\mathbf{z},\boldsymbol{\lambda}\rangle_{(x,t)} = -\langle \mathbf{z}, D^*\boldsymbol{\lambda}\rangle_{(x,t)} + \left\langle \mathbf{z}^L,\left(\mathbf{A}^\top\boldsymbol{\lambda}\right)^L\right\rangle_t - \left\langle \mathbf{z}^0,\left(\mathbf{A}^\top\boldsymbol{\lambda}\right)^0\right\rangle_t \quad (12)$$

Proof: The inner product $\langle D\mathbf{z},\boldsymbol{\lambda}\rangle_{(x,t)}$ in $\mathcal{L}_2$ is

$$\langle D\mathbf{z},\boldsymbol{\lambda}\rangle_{(x,t)} = \int_0^T\int_0^L \boldsymbol{\lambda}^\top D\mathbf{z}\,dxdt$$

Integrating by parts yields

$$\int_0^T\int_0^L \boldsymbol{\lambda}^\top D\mathbf{z}\,dxdt = -\int_0^T\int_0^L \mathbf{z}^\top D^*\boldsymbol{\lambda}\,dxdt + \int_0^T\left[\left(\mathbf{z}^L\right)^\top\left(\mathbf{A}^L\right)^\top\boldsymbol{\lambda}^L - \left(\mathbf{z}^0\right)^\top\left(\mathbf{A}^0\right)^\top\boldsymbol{\lambda}^0\right]dt$$

We define the following inner products

$$\left\langle \mathbf{z}^L,\left(\mathbf{A}^\top\boldsymbol{\lambda}\right)^L\right\rangle_t - \left\langle \mathbf{z}^0,\left(\mathbf{A}^\top\boldsymbol{\lambda}\right)^0\right\rangle_t = \int_0^T\left[\left(\mathbf{z}^L\right)^\top\left(\mathbf{A}^L\right)^\top\boldsymbol{\lambda}^L - \left(\mathbf{z}^0\right)^\top\left(\mathbf{A}^0\right)^\top\boldsymbol{\lambda}^0\right]dt$$

Equation (12) is thus obtained. ■

Definition 1: Let $F: X\to Y$ be a functional with X, Y in Banach spaces and $\boldsymbol{\alpha}\in X$. If there exists a continuous linear operator $\nabla F(\boldsymbol{\alpha}): X\to Y$ for any variation $\boldsymbol{\delta}\in X$ such that

$$\lim_{\varepsilon\to 0}\left\|\nabla F(\boldsymbol{\alpha})\,\boldsymbol{\delta} - \frac{F(\boldsymbol{\alpha}+\varepsilon\boldsymbol{\delta}) - F(\boldsymbol{\alpha})}{\varepsilon}\right\| = 0$$

then $\nabla F(\boldsymbol{\alpha})$ is called a Gâteaux derivative of F at $\boldsymbol{\alpha}$.

We now define the Hamiltonians

$$H_1(\mathbf{y},x,\boldsymbol{\lambda}) = L_1 - \boldsymbol{\lambda}^\top\mathbf{B} \quad (13)$$

$$H_2\left(\mathbf{y}^0,\mathbf{y}^L,\mathbf{u},\mathbf{v},\boldsymbol{\mu}\right) = L_2 + \boldsymbol{\mu}^\top\mathbf{f} \quad (14)$$

We are now ready to state the necessary conditions for optimality.

Theorem 2: If (A1)-(A3) are fulfilled and if $(\bar{\mathbf{y}},\bar{\mathbf{u}},\bar{\mathbf{v}})$ is an optimal solution of (9), then there exist adjoint variables $\boldsymbol{\lambda}(x,t):[0,L]\times[0,T]\to\mathbb{R}^n$ and $\boldsymbol{\mu}:[0,T]\to\mathbb{R}^k$ that satisfy the following dual adjoint system

$$\left.\begin{array}{c}\boldsymbol{\lambda}_t + \left(\mathbf{A}^\top\boldsymbol{\lambda}\right)_x + H_{1,\mathbf{y}}^\top = \mathbf{0}\\ \left(\mathbf{A}^\top\boldsymbol{\lambda}\right)^L = H_{2,\mathbf{y}^L}^\top + \mathbf{g}_{\mathbf{y}^L}^\top\left(\mathbf{A}^\top\boldsymbol{\lambda}\right)^0 + \mathbf{g}_{\mathbf{y}^L}^\top H_{2,\mathbf{y}^0}^\top\\ \boldsymbol{\lambda}(x,T) = \mathbf{0}\end{array}\right\} \quad (15)$$

$$\left.\begin{array}{c}\dot{\boldsymbol{\mu}} = -H_{2,\mathbf{u}}^\top - \mathbf{g}_{\mathbf{u}}^\top\left(\mathbf{A}^\top\boldsymbol{\lambda}\right)^0 - \mathbf{g}_{\mathbf{u}}^\top H_{2,\mathbf{y}^0}^\top\\ \boldsymbol{\mu}(T) = \mathbf{0}\end{array}\right\} \quad (16)$$

with a terminal time transversality condition

$$\int_0^L L_1|_{t=T}\,dx + L_2|_{t=T} = 0 \quad (17)$$

such that the optimal control is one that satisfies the following Pontryagin's minimum principle

$$\bar{\mathbf{v}} = \arg\min_{\mathbf{v}\in\mathcal{V}_{ad}} H_2\left(\bar{\mathbf{y}^0},\bar{\mathbf{y}^L},\bar{\mathbf{u}},\mathbf{v},\boldsymbol{\mu}\right) \quad (18)$$

Proof: Let $\boldsymbol{\alpha} = (\mathbf{z}, \mathbf{p})$ be solutions to $\boldsymbol{\beta} = (\bar{\mathbf{y}}, \bar{\mathbf{u}})$ in variations for a variation $\mathbf{q}$ in $\bar{\mathbf{v}}$, then the variation in the cost functional from (9) is computed as

$$\Delta J(\mathbf{q}) = \nabla J(\boldsymbol{\beta}) + J(\boldsymbol{\beta}, \bar{\mathbf{v}} + \mathbf{q}) - J(\boldsymbol{\beta}, \bar{\mathbf{v}}) > 0$$

where the Gâteaux derivative of J at $\boldsymbol{\beta}$ is evaluated as

$$\begin{aligned}\nabla J(\boldsymbol{\beta}) = &\left\langle H_{1,\mathbf{y}}^{\top}, \mathbf{z}\right\rangle_{(x,t)} - \langle \boldsymbol{\lambda}, \mathbf{z}_t\rangle_{(x,t)} - \langle \boldsymbol{\lambda}, D\mathbf{z}\rangle_{(x,t)} \\ &+ \left\langle H_{2,\mathbf{y}^0}^{\top}, \mathbf{z}^0\right\rangle_t + \left\langle H_{2,\mathbf{y}^L}^{\top}, \mathbf{z}^L\right\rangle_t + \left\langle H_{2,\mathbf{u}}^{\top}, \mathbf{p}\right\rangle_t - \langle \boldsymbol{\mu}, \dot{\mathbf{p}}\rangle_t \\ &+ \delta t \left(\int_0^L L_1|_{t=T}\, dx + L_2|_{t=T}\right)\end{aligned}$$

From the boundary condition (6), we have the following variations

$$\mathbf{z}^0 = \mathbf{g}_{\mathbf{y}^L}^{\top}\mathbf{z}^L + \mathbf{g}_{\mathbf{u}}^{\top}\mathbf{u}$$

From Lemma 1 and the variations in the boundary condition (6) plus vanishing variations in initial conditions for (3) and (8), this becomes

$$\begin{aligned}\nabla J(\boldsymbol{\beta}) = &\left\langle \boldsymbol{\lambda}_t + D^*\boldsymbol{\lambda} + H_{1,\mathbf{y}}^{\top}, \mathbf{z}\right\rangle_{(x,t)} \\ &- \langle \mathbf{z}(x,T), \boldsymbol{\lambda}(x,T)\rangle_x \\ &\left\langle \mathbf{g}_{\mathbf{y}^L}^{\top} H_{2,\mathbf{y}^0}^{\top} + \mathbf{g}_{\mathbf{y}^L}^{\top}\left(\mathbf{A}^{\top}\boldsymbol{\lambda}\right)^0 + H_{2,\mathbf{y}^L}^{\top} - \left(\mathbf{A}^{\top}\boldsymbol{\lambda}\right)^L, \mathbf{z}\right\rangle_t \\ &+ \left\langle \mathbf{g}_{\mathbf{u}}^{\top}\left(\mathbf{A}^{\top}\boldsymbol{\lambda}\right)^0 + \mathbf{g}_{\mathbf{u}}^{\top} H_{2,\mathbf{y}^0}^{\top} + H_{2,\mathbf{u}}^{\top} + \dot{\boldsymbol{\mu}}^{\top}, \mathbf{p}\right\rangle_t \\ &- \boldsymbol{\mu}^{\top}(T)\,\mathbf{p}(T) + \delta t \left(\int_0^L L_1|_{t=T}\, dx + L_2|_{t=T}\right)\end{aligned}$$

Setting $\nabla J(\boldsymbol{\beta}) = 0$ for arbitrary variation $\boldsymbol{\alpha}$ results in (15)-(17). Then the variation in the cost functional becomes

$$\begin{aligned}\Delta J(\mathbf{q}) = &\int_0^T \left[H_2\left(\bar{\mathbf{y}}^0, \bar{\mathbf{y}}^L, \bar{\mathbf{u}}, \bar{\mathbf{v}} + \mathbf{q}, \boldsymbol{\mu}\right) - \boldsymbol{\mu}^{\top}\dot{\bar{u}}\right] dt \\ &- \int_0^T \left[H_2\left(\bar{\mathbf{y}}^0, \bar{\mathbf{y}}^L, \bar{\mathbf{u}}, \bar{\mathbf{v}}, \boldsymbol{\mu}\right) - \boldsymbol{\mu}^{\top}\dot{\bar{u}}\right] dt > 0\end{aligned}$$

This leads to the Pontryagin's minimum principle

$$H_2\left(\bar{\mathbf{y}}^0, \bar{\mathbf{y}}^L, \bar{\mathbf{u}}, \bar{\mathbf{v}} + \mathbf{q}, \boldsymbol{\mu}\right) > H_2\left(\bar{\mathbf{y}}^0, \bar{\mathbf{y}}^L, \bar{\mathbf{u}}, \bar{\mathbf{v}}, \boldsymbol{\mu}\right)$$

for all values of the variation $\mathbf{q}$. ■

Equation (18) is the equivalent Weierstrass condition for strong variations. For weak variations when $\mathbf{v}$ is unconstrained and L_1 and L_2 are convex, the Pontryagin's principle leads to the Legendre-Clebsch condition

$$\left.\begin{aligned}H_{2,\mathbf{v}}\left(\bar{\mathbf{y}}^0, \bar{\mathbf{y}}^L, \bar{\mathbf{u}}, \bar{\mathbf{v}}, \boldsymbol{\mu}\right) &= \mathbf{0} \\ H_{2,\mathbf{vv}}\left(\bar{\mathbf{y}}^0, \bar{\mathbf{y}}^L, \bar{\mathbf{u}}, \bar{\mathbf{v}}, \boldsymbol{\mu}\right) &> \mathbf{0}\end{aligned}\right\} \tag{19}$$

IV. NUMERICAL OPTIMIZATION

Adjoint-based optimization of the coupled hyperbolic-ODE system in the adjoint framework requires solving the solution of the control variable $\mathbf{v}(t)$, which is a function of the adjoint variable $\boldsymbol{\mu}(t)$ according to (19). However, the solutions of the adjoint variables $\boldsymbol{\lambda}(x,t)$ and $\boldsymbol{\mu}(t)$ depend on the solution of the hyperbolic variable $\mathbf{y}(x,t)$ and the boundary control $\mathbf{u}(t)$, which in turn is a function of the control variable $\mathbf{v}(t)$. This circular dependency and the fact that the hyperbolic-ODE system is defined at the initial time while the adjoint dual system is usually specified at the terminal time pose a challenge for the nonlinear constrained optimization.

Thus, for an adjoint-based nonlinear optimization, the solution usually involves a two-point boundary value problem whereby a original hyperbolic-ODE system is solved by forward-time integration from the initial time. On the other hand, the adjoint system is usually solved by backward-time integration from the terminal time when its solution is normally specified. The solution of this optimization problem is further complicated by the nonlinear coupling between the hyperbolic PDE and ODE at the periodic boundary. Since a PDE system is usually solved by discretization of the solution domain into a finite-dimensional system of potentially large number of grid points for improved numerical accuracy, this generally can result in a significantly large-scale optimization problem. The size and complexity of the discretized hyperbolic PDE constraints often pose significant challenges in numerical optimization methods for these systems. For instance, 3-D hyperbolic problems can result in a PDE simulation that can often scale to millions of variables. Large-scale optimization problems are a current subject of research in the field of numerical optimization [15]. Moreover, due to numerical stability, hyperbolic PDE solutions may use explicit numerical scheme that requires very small time steps which can significantly contribute to the computational speed for a single iterative solution.

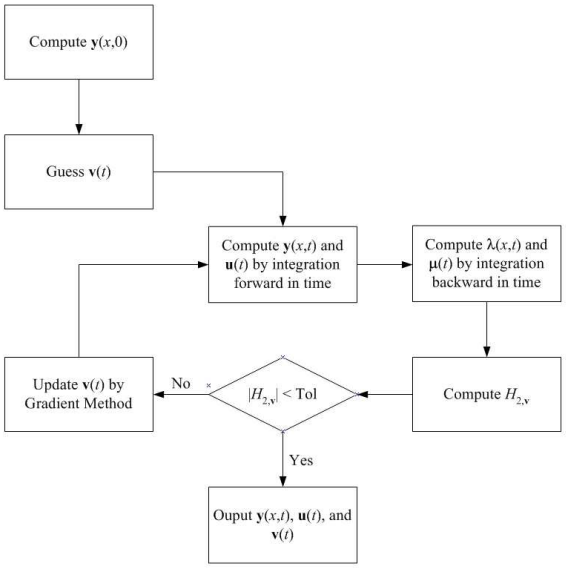

Fig. 2 - Schematic Diagram of Gradient Method

For a general optimization problem, the gradient method has a certain appeal in that the system equations are solved exactly at each iteration, with the control variable being perturbed from step to step to "home in" on the optimal solution. In other words, the algorithm simulates the system dynamic response with varying control histories from one iteration to the next. The solution method usually starts with an initial guess of the control. This then allows the state equations to be completely determined by integration forward in time using the specified

initial conditions. Once the state trajectory is computed, the adjoint equations are integrated backward in time using the teminal time transversality conditions. The control is then updated for the next iteration. The whole solution process is iterated until a convergence on the control is obtained [16]. A schematic of the gradient optimization for this particular PDE-ODE system is shown in Fig. 2.

As seen from Fig. 2, the gradient method utilizes the control gradient of the Hamiltonian function H_2 to compute the control variable $\mathbf{v}$ iteratively according to

$$\mathbf{v}^{(p+1)} = \mathbf{v}^{(p)} - \boldsymbol{\epsilon}^{(p)} H_{2,\mathbf{v}^{(p)}}^{\top} \tag{20}$$

splitted where $\boldsymbol{\epsilon}$ is a positive-definite control-gradient weighting matrix and p is the iteration index.

This update is the well-known steepest-descent algorithm. Visualizing a bowl-shaped Hamiltonian function H_2 as illustrated in Fig. 3, the gradient $H_{2,\mathbf{v}}$ defines the magnitude and direction of the local slope of the Hamiltonian function. Perturbing the control by some function of $H_{2,\mathbf{v}}$ moves the control toward the bottom of the bowl. Proper selection of the step size is critical for a rapid convergence to the minimizing control. If $\boldsymbol{\epsilon}$ is too small, the convergence may require a large number of iterations. On the other hand, if $\boldsymbol{\epsilon}$ is too large, the convergence may not occur at all. Therefore, the effectiveness of a gradient method is predicated upon a judicious choice of the weighting matrix $\boldsymbol{\epsilon}$. If the weighting matrix $\boldsymbol{\epsilon}$ is the inverse of the Hessian matrix of the Hamiltonian function H_2, then the method is known as a second-order gradient or Newton-Raphson method as [17]

$$\mathbf{v}^{(p+1)} = \mathbf{v}^{(p)} - \left[H_{2,\mathbf{vv}}^{(p)}\right]^{-1} H_{2,\mathbf{v}^{(p)}}^{\top} \tag{21}$$

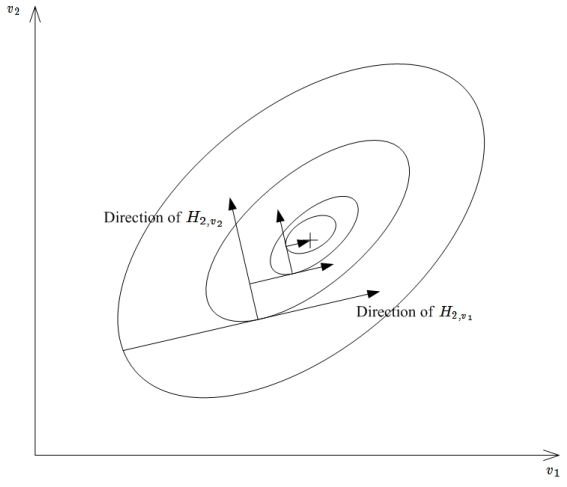

Fig. 3 - Steepest Descent Approach to a Minimum Hamiltonian Function

If H_2 is quadratic in $\mathbf{v}$, then the method should converge within two iterations [17]. However, if the initial guess to start the method is too far off the actual solution, the method may not converge at all. In practice, the Hessian matrix $H_{2,\mathbf{vv}}$ is time consuming to calculate analytically, but a numerical evaluation of the Hessian matrix may be subject to a numerical accuracy problem.

V. COMPUTATIONAL METHODS

To implement the gradient algorithm, we need to solve the hyperbolic and adjoint systems. The simplest type of spatial discretization for a first-order, hyperbolic PDE is the first-order upwind finite-difference method [18]. In flow with shocks, spatial discretization may not have sufficient stability requirements and could cause numerical oscillations at the discontinuities. Various spatial discretization schemes such as the Lax-Wendroff method incorporate an artificial viscosity to dampen these numerical oscillations. From the physical interpretation of wave speeds, the eigenvalues of the matrix $\mathbf{A}$ dictates the type of numerical methods for the hyperbolic PDEs. The most general case is when the matrix $\mathbf{A}$ has mixed-sign eigenvalues. In this case, the matrix $\mathbf{A}$ can be splitted into a semi-positive definite matrix and a semi-negative definite matrix as

$$\mathbf{A} = \mathbf{A}^{+} + \mathbf{A}^{-} \tag{22}$$

where $\mathbf{A}^{+} = \boldsymbol{\Phi}\boldsymbol{\Lambda}^{+}\boldsymbol{\Phi}^{-1}$ and $\mathbf{A}^{-} = \boldsymbol{\Phi}\boldsymbol{\Lambda}^{-}\boldsymbol{\Phi}^{-1}$.

Equation (3) can be written in a wave-splitting form as

$$\boldsymbol{\Phi}^{-1}\mathbf{y}_t + \boldsymbol{\Lambda}^{+}\boldsymbol{\Phi}^{-1}\mathbf{y}_x + \boldsymbol{\Lambda}^{-}\boldsymbol{\Phi}^{-1}\mathbf{y}_x + \boldsymbol{\Phi}^{-1}\mathbf{B} = \mathbf{0} \tag{23}$$

We let $\boldsymbol{\Phi}^{-1} = \boldsymbol{\Psi}^{+} + \boldsymbol{\Psi}^{-}$ where $\boldsymbol{\Psi}^{+} = \boldsymbol{\Lambda}^{+}\boldsymbol{\Lambda}^{-1}\boldsymbol{\Phi}^{-1}$ and $\boldsymbol{\Psi}^{-} = \boldsymbol{\Lambda}^{-}\boldsymbol{\Lambda}^{-1}\boldsymbol{\Phi}^{-1}$. We note that if $\boldsymbol{\Lambda} = diag\left(\lambda_1, \ldots, \lambda_m, \lambda_{m+1}, \ldots, \lambda_n\right)$ where $\lambda_i < 0$ for $i = 1, \ldots, m$ and $\lambda_i > 0$ for $i = m+1, \ldots, n$, then $\boldsymbol{\Lambda}^{+}\boldsymbol{\Lambda}^{-1} = diag\left(0, \ldots, 0, 1, \ldots, 1\right)$ and $\boldsymbol{\Lambda}^{-}\boldsymbol{\Lambda}^{-1} = diag\left(1, \ldots, 1, 0, \ldots, 0\right)$. In order for (23) to converge to a correct steady state solution, we need to maintain a numerical consistency between the steady state form and the hyperbolic form. This is done by evaluating all the matrices in (23) at a prior grid point for the positive wave speeds and at a current grid point for the negative wave speed. Using this approach, (23) is now discretized using a wave-splitting, first-order upwind finite-difference method as

$$\boldsymbol{\Psi}_{i-1}^{+}\frac{d\mathbf{y}_i}{dt} + \boldsymbol{\Lambda}_{i-1}^{+}\boldsymbol{\Psi}_{i-1}^{+}\frac{\mathbf{y}_i - \mathbf{y}_{i-1}}{\Delta x} + \boldsymbol{\Psi}_{i-1}^{+}\mathbf{B}_{i-1} = 0 \tag{24}$$

$$\boldsymbol{\Psi}_{i}^{-}\frac{d\mathbf{y}_i}{dt} + \boldsymbol{\Lambda}_{i}^{-}\boldsymbol{\Psi}_{i}^{-}\frac{\mathbf{y}_{i+1} - \mathbf{y}_i}{\Delta x} + \boldsymbol{\Psi}_{i}^{-}\mathbf{B}_{i} = 0 \tag{25}$$

where $i = 2, 3, \ldots, M-1$ denotes the index of the interior points not on the boundary such that $x_i = \frac{i-1}{M-1}L$ and $\mathbf{y}_i\left(t\right) = \mathbf{y}\left(x_i, t\right)$.

We now combine Eqs. (24) to (25) into a vector form as

$$\frac{d\mathbf{y}_i}{dt} + \mathbf{A}_{i-1}^{+}\frac{\mathbf{y}_i - \mathbf{y}_{i-1}}{\Delta x} + \mathbf{A}_{i}^{-}\frac{\mathbf{y}_{i+1} - \mathbf{y}_i}{\Delta x} + \mathbf{B}_{i-1/2} = \mathbf{0} \tag{26}$$

where

$$\mathbf{A}_{i-1}^{+} = \left(\boldsymbol{\Psi}_{i-1}^{+} + \boldsymbol{\Psi}_{i}^{-}\right)^{-1}\boldsymbol{\Lambda}_{i-1}^{+}\boldsymbol{\Psi}_{i-1}^{+}$$

$$\mathbf{A}_{i}^{-} = \left(\boldsymbol{\Psi}_{i-1}^{+} + \boldsymbol{\Psi}_{i}^{-}\right)^{-1}\boldsymbol{\Lambda}_{i}^{-}\boldsymbol{\Psi}_{i}^{-}$$

$$\mathbf{B}_{i-1/2} = \left(\boldsymbol{\Psi}_{i-1}^{+} + \boldsymbol{\Psi}_{i}^{-}\right)^{-1}\left(\boldsymbol{\Psi}_{i-1}^{+}\mathbf{B}_{i-1} + \boldsymbol{\Psi}_{i}^{-}\mathbf{B}_{i}\right)$$

To solve for Eq. (26), we need the information on the system boundaries imposed by the periodic boundary condition (6). Since $\mathbf{g}_{\mathbf{y}(L,t)}$ is full rank, $\mathbf{g}$ is invertible with respect to

$\mathbf{y}(L,t)$. Therefore, the periodic boundary condition (6) is equivalent to

$$\mathbf{y}(L,t) = \mathbf{g}^{-1}(\mathbf{y}(0,t), \mathbf{u}) \tag{27}$$

The nonlinear periodic boundary condition (6) can now be decomposed into

$$\mathbf{\Lambda}^{+}\mathbf{\Lambda}^{-1}\mathbf{y}_{1,j+1} = \mathbf{G}(\mathbf{y}_{M,j+1}, \mathbf{u}_{j+1}) \tag{28}$$

$$\mathbf{\Lambda}^{-}\mathbf{\Lambda}^{-1}\mathbf{y}_{M,j+1} = \mathbf{H}(\mathbf{y}_{1,j+1}, \mathbf{u}_{j+1}) \tag{29}$$

where $\mathbf{G} = \mathbf{\Lambda}^{+}\mathbf{\Lambda}^{-1}\mathbf{g}$ and $\mathbf{H} = \mathbf{\Lambda}^{-}\mathbf{\Lambda}^{-1}\mathbf{g}^{-1}$, and $j = 1, 2, \ldots, N-1$ denotes the time index such that $t_j = \frac{j-1}{N-1}T$.

At the incoming boundary $x = 0$, $i = 1$ so that (**??**) cannot admit a positive wave which would require the solution to include a point upstream of $x = 0$ that is nonexistent. Therefore, only the negative wave speed characteristic equation (25) is admitted. Using the Euler's method, we combine (25) with (28) as

$$\left(\mathbf{\Psi}^{-}_{1,j} + \mathbf{\Lambda}^{+}\mathbf{\Lambda}^{-1}\right)\mathbf{y}_{1,j+1} - \mathbf{G}(\mathbf{y}_{M,j+1}, \mathbf{u}_{j+1}) = \mathbf{\Psi}^{-}_{1,j}\mathbf{y}_{1,j} - \mathbf{\Lambda}^{-}_{i,j}\mathbf{\Psi}^{-}_{i,j}\left(\frac{\mathbf{y}_{2,j} - \mathbf{y}_{1,j}}{\Delta x}\right)\Delta t - \mathbf{\Psi}^{-}_{i,j}\mathbf{B}_{i,j}\Delta t \tag{30}$$

where the time step Δt is chosen to satisfy the following condition

$$\Delta t \leq \min\left\{\frac{\Delta x}{(1+\alpha)\max|\lambda(\mathbf{A})|}, \frac{2}{\max|\lambda(\mathbf{f_u})|}\right\} \tag{31}$$

The first condition is the Courant-Friedrichs-Levy (CFL) condition for (26). The term $1+\alpha$ is due to the contribution of the source term $\mathbf{B}$. The second condition is the numerical stability condition for (8) using an explicit time integration scheme.

We next consider the boundary at $x = L$, $i = M$. The situation is now reverse whereby the solution can only admit the positive eigenvalue characteristic equation (24) which can be combined with (29) to yield

$$\left(\mathbf{\Psi}^{+}_{m-1,j} + \mathbf{\Lambda}^{-}\mathbf{\Lambda}^{-1}\right)\mathbf{y}_{m,j+1} - \mathbf{H}(\mathbf{y}_{1,j+1}, \mathbf{u}_{j+1}) = \mathbf{\Psi}^{+}_{m-1,j}\mathbf{y}_{mj} - \mathbf{\Lambda}^{+}_{m-1,j}\mathbf{\Psi}^{+}_{m-1,j}\left(\frac{\mathbf{y}_{m,j} - \mathbf{y}_{m-1,j}}{\Delta x}\right)\Delta t - \mathbf{\Psi}^{+}_{m-1,j}\mathbf{B}_{m-1,j}\Delta t \tag{32}$$

Equations (30) and (32) are nonlinear. An iterative method is implemented to search for the zero solution of $\mathbf{y}_{1,j+1}$ and $\mathbf{y}_{m,j+1}$ using information from the previous time step. To solve these equations, the boundary control $\mathbf{u}$ must be determined for a given time history of the control $\mathbf{v}$ by integrating (8). Once the information at the system boundary has been determined, the solution of $\mathbf{y}_{i,j+1}$ at the interior points can be found by integrating (26) using Runge-Kutta or Euler's method.

In order to compute the adjoint dual system, we transform (15) and (16) by reversing the space and time variables using a distance-to-go variable $\chi = L - x$ and a time-to-go variable $\tau = T - t$. We also let $\boldsymbol{\vartheta}(\chi,\tau) = \boldsymbol{\lambda}(x,t)$ and $\boldsymbol{\nu}(\tau) = \boldsymbol{\mu}(t)$. Then, (15) and (16) becoms

$$\boldsymbol{\vartheta}_{\tau} + \left(\mathbf{A}^{\top}\boldsymbol{\vartheta}\right)_{\chi} + \mathbf{B}^{\top}_{\mathbf{y}}\boldsymbol{\vartheta} - L^{\top}_{1,\mathbf{y}} = \mathbf{0} \tag{33}$$

$$\left(\mathbf{A}^{\top}\right)^{L}\boldsymbol{\vartheta}^{0} = \mathbf{g}^{\top}_{\mathbf{y}^{L}}\left(\mathbf{A}^{\top}\right)^{0}\boldsymbol{\vartheta}^{L} + \left(\mathbf{f}^{\top}_{\mathbf{y}^{L}} + \mathbf{g}^{\top}_{\mathbf{y}^{L}}\mathbf{f}^{\top}_{\mathbf{y}^{0}}\right)\boldsymbol{\nu} \tag{34}$$

$$\dot{\boldsymbol{\nu}} = \left(\mathbf{f}^{\top}_{\mathbf{u}} + \mathbf{g}^{\top}_{\mathbf{u}}\mathbf{f}^{\top}_{\mathbf{y}^{0}}\right)\boldsymbol{\nu} + \mathbf{g}^{\top}_{\mathbf{u}}\left(\mathbf{A}^{\top}\right)^{0}\boldsymbol{\vartheta}^{L} + L^{\top}_{2,\mathbf{u}} \tag{35}$$

The advantage of the reversal transformation is that it generally results in a stable adjoint equation backward in space and time when the hyperbolic equation is normally stable forward in time. This stability is necessary for any numerical optimization process to succeed. Equation (33) can be discretized as

$$\frac{d\boldsymbol{\vartheta}_k}{d\tau} + \frac{\left(\mathbf{A}^{+}_{i}\right)^{\top}\boldsymbol{\vartheta}_k - \left(\mathbf{A}^{+}_{i+1}\right)^{\top}\boldsymbol{\vartheta}_{k-1}}{\Delta\chi} + \frac{\left(\mathbf{A}^{-}_{i-1}\right)^{\top}\boldsymbol{\vartheta}_{k+1} - \left(\mathbf{A}^{-}_{i}\right)^{\top}\boldsymbol{\vartheta}_k}{\Delta\chi} + \mathbf{\Lambda}^{+}\mathbf{\Lambda}^{-1}\left(\mathbf{B}^{\top}_{\mathbf{y}_{i+1}}\boldsymbol{\vartheta}_{k-1} - L^{\top}_{1,\mathbf{y}_{i+1}}\right) + \mathbf{\Lambda}^{-}\mathbf{\Lambda}^{-1}\left(\mathbf{B}^{\top}_{\mathbf{y}_i}\boldsymbol{\vartheta}_k - L^{\top}_{1,\mathbf{y}_i}\right) = \mathbf{0} \tag{36}$$

where $k = 2, 3, \ldots, M-1$ represents the index of χ_k such that $k = M - i + 1$.

Generally, adjoint equations are quite sensitive to numerical round-off errors and differences in time scales, which may potentially cause numerical instability. In order to prevent numerical errors, an implicit scheme for integrating (36) could be used, but the numerical method is much more complex than an explicit scheme and involve a full matrix inversion at each time step. The implicit-scheme discretization is

$$\frac{\boldsymbol{\vartheta}_{k,l+1} - \boldsymbol{\vartheta}_{k,l}}{\Delta\tau} + \frac{\left(\mathbf{A}^{+}_{i,j-1}\right)^{\top}\boldsymbol{\vartheta}_{k,l+1} - \left(\mathbf{A}^{+}_{i+1,j-1}\right)^{\top}\boldsymbol{\vartheta}_{k-1,l+1}}{\Delta\chi} + \frac{\left(\mathbf{A}^{-}_{i-1,j-1}\right)^{\top}\boldsymbol{\vartheta}_{k+1,l+1} - \left(\mathbf{A}^{-}_{i,j-1}\right)^{\top}\boldsymbol{\vartheta}_{k,l+1}}{\Delta\chi} + \mathbf{\Lambda}^{-}\mathbf{\Lambda}^{-1}\left(\mathbf{B}^{\top}_{\mathbf{y}_{i+1,j-1}}\boldsymbol{\vartheta}_{k-1,l+1} - L^{\top}_{1,\mathbf{y}_{i+1,j-1}}\right) + \mathbf{\Lambda}^{-}\mathbf{\Lambda}^{-1}\left(\mathbf{B}^{\top}_{\mathbf{y}_{i,j-1}}\boldsymbol{\vartheta}_{k,l+1} - L^{\top}_{1,\mathbf{y}_{i,j-1}}\right) = \mathbf{0} \tag{37}$$

where $l = 2, 3, \ldots, N$ represents the index of τ_l such that $l = N - j + 1$.

At the boundary, the adjoint solution is treated in a similar manner as before whereby the characteristic adjoint equations for the positive and negative wave speeds are separated along with the boundary condition (34)

$$\frac{\boldsymbol{\vartheta}_{1,l+1} - \boldsymbol{\vartheta}_{1,l}}{\Delta\tau} + \frac{\left(\mathbf{A}^{-}_{M-1,j-1}\right)^{\top}\boldsymbol{\vartheta}_{2,l+1} - \left(\mathbf{A}^{-}_{M,j-1}\right)^{\top}\boldsymbol{\vartheta}_{1,l+1}}{\Delta\chi} + \mathbf{\Lambda}^{-}\mathbf{\Lambda}^{-1}\left(\mathbf{B}^{\top}_{\mathbf{y}_{M,j-1}}\boldsymbol{\vartheta}_{1,l+1} - L^{\top}_{1,\mathbf{y}_{M,j-1}}\right) = \mathbf{0} \tag{38}$$

$$\frac{\boldsymbol{\vartheta}_{M,l+1} - \boldsymbol{\vartheta}_{M,l}}{\Delta\tau} + \frac{\left(\mathbf{A}^{+}_{1,j-1}\right)^{\top}\boldsymbol{\vartheta}_{M,l+1} - \left(\mathbf{A}^{+}_{2,j-1}\right)^{\top}\boldsymbol{\vartheta}_{M-1,l+1}}{\Delta\chi} + \mathbf{\Lambda}^{+}\mathbf{\Lambda}^{-1}\left(\mathbf{B}^{\top}_{\mathbf{y}_{2,j-1}}\boldsymbol{\vartheta}_{M-1,l+1} - L^{\top}_{1,\mathbf{y}_{2,j-1}}\right) = \mathbf{0} \tag{39}$$

$$\boldsymbol{\vartheta}_{1,l+1} = \mathbf{A}^{-\top}_{M,j-1}\mathbf{g}^{\top}_{\mathbf{y}_{M,j-1}}\mathbf{A}^{\top}_{1,j-1}\boldsymbol{\vartheta}_{M,l+1} + \mathbf{A}^{-\top}_{M,j-1}\left(\mathbf{f}^{\top}_{\mathbf{y}_{M,j-1}} + \mathbf{g}^{\top}_{\mathbf{y}_{M,j-1}}\mathbf{f}^{\top}_{\mathbf{y}_{1,j-1}}\right)\boldsymbol{\nu}_{l+1} \tag{40}$$

Equations (37) to (40) must be assembled into a linear matrix equation for a simultaneous solution at a future time step $l+1$ as a function of the current time step l.

Another approach to improving the numerical stability of the adjoint equation (33) is by treating it as a quasi-steady state approximation. This is possible if the time scale is substantially different between the adjoint PDE (33) and the adjoint ODE (35). This time scale difference can be inferred by the magnitude of the largest eigenvalues of $\mathbf{f_u}$ and $\mathbf{A}$. Generally, transport systems carry information by wave propagation at a faster speed than the dynamics imposed at the system boundaries. Another justification for this approach is that the control variable $\mathbf{v}(t)$ that optimizes the hyperbolic system is a direct function of the adjoint variable $\boldsymbol{\mu}(t)$, but is only indirectly influenced by the adjoint variable $\boldsymbol{\lambda}(x,t)$ on the boundary $x = 0$. Therefore, the quasi-steady state approximation for the adjoint equation (33) is reasonable.

Let $\boldsymbol{z}(\chi,\tau) = \mathbf{A}^T(\mathbf{y},x)\boldsymbol{\lambda}(x,t)$, then, (15) and (16) become

$$\boldsymbol{z}_\chi + \mathbf{B}_\mathbf{y}^\top \mathbf{A}^{-\top}\boldsymbol{z} - L_{1,\mathbf{y}}^\top = \mathbf{0} \tag{41}$$

$$\boldsymbol{z}^0 = \mathbf{g}_{\mathbf{y}^L}^\top \boldsymbol{z}^L + \left(\mathbf{f}_{\mathbf{y}^L}^\top + \mathbf{g}_{\mathbf{y}^L}^\top \mathbf{f}_{\mathbf{y}^0}^\top\right)\boldsymbol{\nu} \tag{42}$$

$$\dot{\boldsymbol{\nu}} = \left(\mathbf{f}_\mathbf{u}^\top + \mathbf{g}_\mathbf{u}^\top \mathbf{f}_{\mathbf{y}^0}^\top\right)\boldsymbol{\nu} + \mathbf{g}_\mathbf{u}^\top \boldsymbol{z}^L + L_{2,\mathbf{u}}^\top \tag{43}$$

A first order finite-difference discretization can be used to solve (41) and the boundary condition

$$\boldsymbol{z}_{k+1} = \boldsymbol{z}_k - \Delta\chi\left(\mathbf{B}_{\mathbf{y},i}^\top \mathbf{A}_i^{-\top}\boldsymbol{z}_k + L_{1,\mathbf{y}_i}^\top\right) \tag{44}$$

$$\boldsymbol{z}_1 = \mathbf{g}_{\mathbf{y}_M}^\top \boldsymbol{z}_M + \left(\mathbf{f}_{\mathbf{y}_M}^\top + \mathbf{g}_{\mathbf{y}_M}^\top \mathbf{f}_{\mathbf{y}_1}^\top\right)\boldsymbol{\nu} \tag{45}$$

Since the boundary condition (42) is periodic, the solution of $\boldsymbol{z}$ results in a full matrix inversion at each time step. In addition, $\boldsymbol{\nu}$ must be solved simultaneously from (43) at each time step. Comparing this to the implicit scheme, the quasi-steady state approach requires fewer operations and hence a more efficient method for solving the adjoint equation (15).

Based on the second-order gradient method, each each iteration, the control variable $\mathbf{v}$ is updated as

$$\mathbf{v}^{(p+1)} = \mathbf{v}^{(p)} - \left[L_{2,\mathbf{v}\mathbf{v}^{(p)}}^{(p)}\right]^{-1}\left(\mathbf{f}_\mathbf{v}^\top \boldsymbol{\mu} + L_{2,\mathbf{v}^{(p)}}^\top\right) \tag{46}$$

where we have replaced $H_{2,\mathbf{vv}}$ with $L_{2,\mathbf{vv}}$ by ignoring the term $\mathbf{f_{vv}}\boldsymbol{\mu}$ in order to simplify the calculation.

Thus, the control variable $\mathbf{v}$ depends on the solution of $\boldsymbol{\mu}$ obtained from $\boldsymbol{\nu}$ which requires solving the adjoint PDE and ODE simultaneously. Thus, initially we guess a time history for the control variable $\mathbf{v}(t)$. This information is then used to solve the hyperbolic equation for $\mathbf{y}(x,t)$ forward in time. The complete time history of $\mathbf{y}(x,t)$ based on the initial guess of the control $\mathbf{v}(t)$ is then used to solve the adjoint equations backward in time to obtain $\boldsymbol{\lambda}(x,t)$ and $\boldsymbol{\mu}(t)$ simultaneously. The control variable $\mathbf{v}(t)$ is then updated in the next iteration from (46). This process is repeated until the solution of the control $\mathbf{v}(t)$ converges.

VI. NUMERICAL EXAMPLE

We now demonstrate the general theory by a numerical example of a closed-loop fluid transport problem in a closed-circuit wind tunnel as shown in Fig. 4.

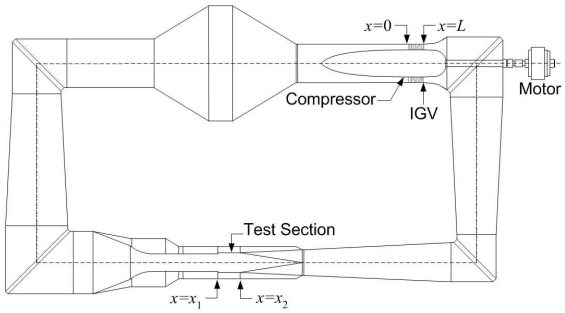

Fig. 4 - Closed-Circuit Wind Tunnel

The fluid transport process in a closed-circuit wind tunnel is a good example of a closed-loop transport process whereby the fluid flow is recirculated through a compressor providing the boundary control action. The compressor is controlled by two auxiliary dynamical processes: a drive motor dynamics, and an inlet guide vane (IGV) dynamics. By controlling the drive motor speed and the IGV angle, the flow in the test section can be set as desired. The fluid flow is governed by the following hyperbolic system with a set of conserved variables; namely, mass flow $\dot{m}$, total pressure p_0, and total temperature T_0. We let $\mathbf{y} = \begin{bmatrix} \dot{m} & p_0 & T_0 \end{bmatrix}$, then according to (3), the matrices $\mathbf{A}$ and $\mathbf{B}$ are defined as

$$\mathbf{A} = \begin{bmatrix} u & \frac{pA}{p_0} & \frac{\dot{m}u}{2T_0} \\ \frac{\rho_0 c^2}{\rho A} & u\left[1-\frac{(\gamma-1)T}{T_0}\right] & \frac{\rho_0 c^2 u}{T_0} \\ \frac{(\gamma-1)T}{\rho A} & -\frac{(\gamma-1)^2 uT}{kp_0} & u\left[1+\frac{(\gamma-1)T}{T_0}\right] \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} \frac{\dot{m}u}{2}\xi \\ \frac{\rho_0 u^3}{2}\left(\frac{T_0}{T}-\gamma+1\right)\xi \\ -(\gamma-1)uT\xi\left(\frac{T_0}{T}-1\right) \end{bmatrix}$$

where u is the flow speed, p is the static pressure, T is the static temperature, $\rho$is the static density, c is the speed of sound, M is the Mach number, A is the flow area, ξ is the loss factor, and γ is the specific heat ratio.

The compressor provides a periodic boundary control at $x = 0$ and $x = L$ which relates the conservation of mass flow, the total pressure rise, and the total temperature rise that enable the flow recirculation as a function of the boundary control $\mathbf{u}$ that comprises the compressor speed ω and the IGV blade angle θ. The drive motor and IGV dynamics dictate constraints on the compressor speed and the IGV blade angle in such a way that these parameters must obey the time varying dynamics of these systems according to (8), where $\mathbf{f}$ represents the dynamical equation and $\mathbf{v}$ represents the control parameters of the drive motor and IGV systems which in this case are the drive motor rotor resistance and the IGV voltage.

A typical wind tunnel operation is to transfer the flow condition in the test section from one Mach number to another. Fig. 5 shows a typical operating envelope for a wind tunnel. What we are interested in is to search for an optimal path of the

Mach number transfer between two fix compressor operating speeds as shown in Fig. 5 that minimizes the cost functional defined in (9) with

$$L_1 = \frac{1}{2}\left(\mathbf{y}-\mathbf{y}_d\right)^\top \mathbf{P}\left(\mathbf{y}-\mathbf{y}_d\right)$$

$$L_2 = \frac{1}{2}\left(\mathbf{u}-\mathbf{u}_d\right)^\top \mathbf{Q}\left(\mathbf{u}-\mathbf{u}_d\right) + \frac{1}{2}\left(\mathbf{v}-\mathbf{v}_d\right)^\top \mathbf{R}\left(\mathbf{v}-\mathbf{v}_d\right)$$

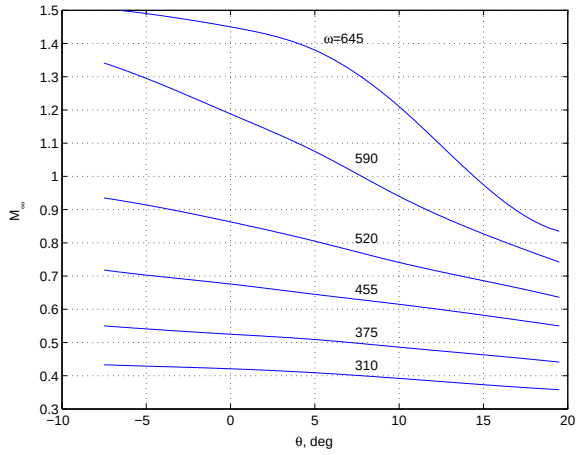

Fig. 5 - Wind Tunnel Operating Envelope

The objective functions L_1 and L_2 are convex for some weighting matrices $\mathbf{P} \geq \mathbf{0}$, $\mathbf{Q} \geq \mathbf{0}$, and $\mathbf{R} > \mathbf{0}$ and are designed to seek a desired terminal condition, denoted by the subscript d, from some initial condition. The optimization begins with an initial guess of the control variable $\mathbf{v}$ which represents the control parameters for the drive motor and IGV systems. The flow solution in the wind tunnel is then computed forward in space and time from the method developed in Section 5. Using the computed time histories of all the flow variables, the adjoint solution is computed backward in space and time using the implicit scheme first and then the quasi-steady state scheme. A minor convergence problem with the implicit scheme is observed when the solution appears to have reached the bottom of the gradient "bowl". The cost increases slightly instead of remaining stable. On the other hand, the quasi-steady state scheme demonstrates a good convergence behavior. The convergence of the optimization is rapid after only four iterations.

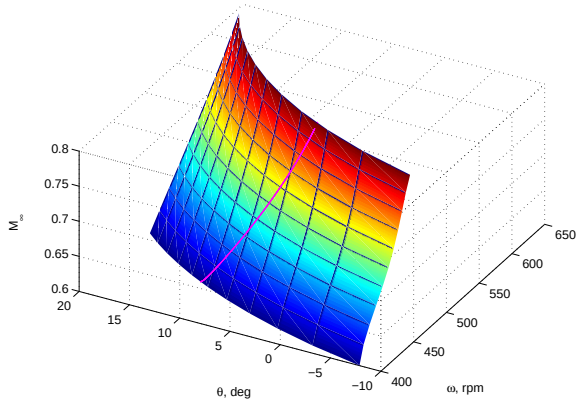

Fig. 6 - Optimal Path Transition

Fig. 6 shows the optimal path of the Mach number transfer. Fig. 7 illustrates the convergence of the control variable $\mathbf{v}$ after just four iterations. Fig. 8 is a plot of the value of the cost functional driven close to a minimum by the gradient method. Fig. 9 is a plot of the hyperbolic solution that shows the variation of Mach number in the flow as a function of space and time.

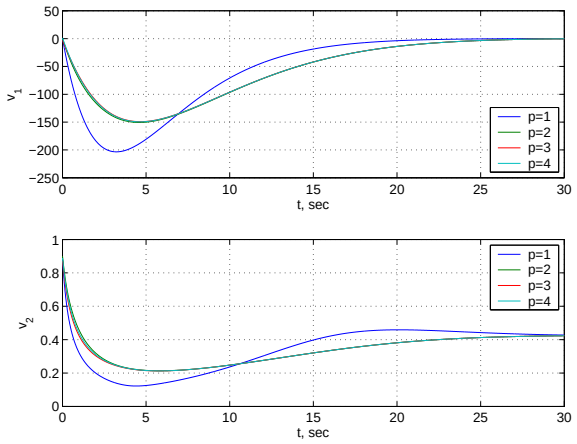

Fig. 7 - Optimal Control Variable

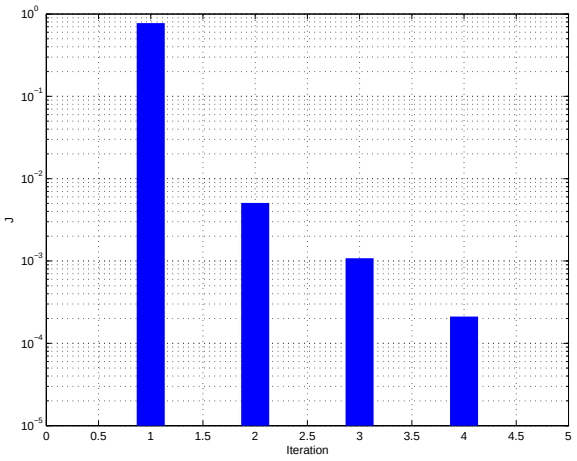

Fig. 8 - Cost Functional Value

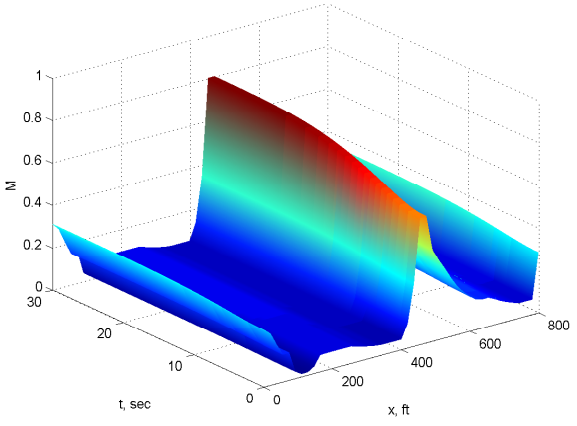

Fig. 9 - Hyperbolic Solution of Mach Number

VII. CONCLUSIONS

We have presented an indirect numerical optimization method using a continuous adjoint approach for a hyperbolic equation with a nonlinear periodic boundary condition subject to a nonlinear differential equation constraints. The continuous adjoint-based optimization method provides a certain advantage in that the adjoint solution can be formulated using a different discretization scheme since the adjoint equation is also a hyperbolic equation. A dual adjoint system is formulated for a multi-objective cost functional comprising an objective function defined over the hyperbolic solution domain and a time-based objective function. The numerical method is based on a second-order gradient method that can provide an accelerated convergence if an initial starting solution is judiciously selected. The hyperbolic equation is solved in forward space and forward time from the initial condition and the periodic boundary condition sing an explicit scheme, wave splitting method. Because of numerical stability, the adjoint hyperbolic equation is integrated backward in time using an implicit scheme. A quasi-steady state method for solving the adjoint equation is introduced. A numerical example is performed for a closed-loop fluid transport system. The results show a fast convergence of the second-order gradient method. The quasi-steady state method results in a more stable solution of the adjoint equation than the implicit scheme.

REFERENCES

[1] Debnath, L., *Nonlinear Partial Differential Equations for Scientists and Engineers*, Birkhauser, Boston, 1997.

[2] Rothfarb, B., Frank, H., Rosenbaum, D. M., Steiglitz, K., Kleitman, D. J., "Optimal Design of Offshore Natural-Gas Pipeline Systems", Journal of the Operations Research Society of America, Vol. 18, No. 6, December 1970, pp. 992-1020.

[3] Bayen, A. M., Raffard, R., Tomlin, C. J., "Adjoint-Based Constrained Control of Eulerian Transportation Networks: Application to Air Traffic Control", Proceeding of the 2004 American Control Conference, June 2004, pp. 5539-5545.

[4] Lee, H. Y., Lee, H.-W., Kim, D., "Dynamic States of a Continuum Traffic Equation with On-Ramp", Physical Review E, Vol. 59, No. 5, May 1999, pp. 5101-5111.

[5] J. F. Wendt, *Computational Fluid Dynamics - An Introduction*, 2nd Edition, Springer-Verlag, 1996.

[6] P. C. Shields, *Elementary Linear Algebra*, Worth Publishers, Inc., New York, 1968.

[7] Diller, D., "A Note on the Uniqueness of Entropy Solutions to First Order Quasilinear Equations", Electronic Journal of Differential Equations, Vol. 1994, No. 5, 1994, pp. 1-4.

[8] Hou, L. S., and Yan, Y., "Dynamics for Controlled Navier-Stokes Systems with Distributed Controls", SIAM Journal on Control and Optimization, Vol. 35, No. 2, 1997, pp. 654-677.

[9] Nguyen, N., Bright, M., and Culley, D., "Feedback Interior Pointwise Control of Euler Equations for Flow Separation Control in Compressors", AIAA Flow Control Conference, AIAA-2006-3020, June 2006.

[10] Raymond, J. and Zidani, H., "Pontryagin's Principle for State-Constrained Control Problems Governed by Parabolic Equations with Unbounded Controls", SIAM Journal on Control and Optimization, Vol. 36, No. 6, 1998, pp. 1853-1879.

[11] Casas, E., Tröltzsch, F., and Unger, A., "Second Order Necessary Optimality Conditions for some State Constrained Control Problems of Semilinear Elliptic Equations", SIAM Journal on Control and Optimization, Vol. 39, 1999, pp. 211-228.

[12] Kazemi, M. A., "A Gradient Technique for an Optimal Control Problem Governed by a System of Nonlinear First Order Partial Differential Equations", Journal of Australian Mathematical Society, Ser. B 36, 1994, 261-273.

[13] Jameson, A., Pierce, N., Martinelli, L., "Optimum Aerodynamic Design Using the Navier-Stokes Equations", Theoretical Computational Fluid Dynamics, Vol. 10, 1998, pp. 213-237.

[14] Nadarajah, S., and Jameson, A., "A Comparison of The Continuous and Discrete Adjoint Approach to Automatic Aerodynamic Optimization", AIAA Conference, AIAA-2000-0667, January 2000.

[15] D. W. Hearn, W.W. Hager, P. M. Pardalos, *Large-Scale Optimization: State of the Art*, Kluwer Academic Publishers, New York, 1994.

[16] D. A. Pierce, *Optimization Theory with Applications*, Dover Publications, New York, 1986.

[17] R. F. Stengel, *Optimal Control and Estimation*, Dover Publications, New York, 1994.

[18] Hirsh, C., *Numerical Computation of Internal and External Flows*, Vol. 1, John Wiley & Sons, Brussels, 1991.